ALGEBRA

LECTURES DELIVERED TO POST-GRADUATE STUDENTS OF CALCUTTA UNIVERSITY

BY

FRIEDRICH WILHELM LEVI, DR. PHIL. NAT.
HARDINGE PROFESSOR

PART I

SYSTEMS OF LINEAR EQUATIONS

PUBLISHED BY THE
UNIVERSITY OF CALCUTTA
1936

BCU 1789

PRINTED AND PUBLISHED BY BHUPENDRALAL BANERJEE
AT THE CALCUTTA UNIVERSITY PRESS, SENATE HOUSE, CALCUTTA.

Reg. No. 988B—August, 1936—E

PREFACE

On introducing a new course of lectures in Algebra I realized after delivering a few lectures that the students of this country should have in their hands a book covering the whole subject-matter of the lectures. In order to make my lectures successful I had no other alternative than to write a text-book and to publish it in different parts as quickly as possible.

So a provisory edition of this text-book is taken in hand. References to the original papers, examples, explanation of details, everything that causes delay of publication had to be omitted in these "lectures." Later on a full text-book on Algebra will be published.

In placing this fascicule in the hands of the students, I offer my heartiest thanks to our energetic Vice-Chancellor, SYAMAPRASAD MOOKERJEE, Esq., M.A., B.L., Barrister-at-Law, M.L.C., without whose sympathetic co-operation this publication would not have come into being. I thank also the Calcutta University Press for having printed this paper in a very short time under difficult circumstances.

Proofs and manuscripts have been revised by Mr. R. C. Bose, M.A., Mr. S. K. Bhar, M.Sc., and especially by Mr. A. C. Choudhury, M.Sc. If the reader do not find many offences against the spirit of the English language, he should be thankful to these three young colleagues of mine.

CALCUTTA, ASUTOSH BUILDING.
August, 1936.

F. W. LEVI.

§ 1. INTRODUCTION.

The problem we have to deal with in this first part is the following:

Given $m(n+1)$ arbitrary numbers

$$a_0, a_1, a_2, \ldots\ldots\ldots, a_n.$$
$$b_0, b_1, b_2, \ldots\ldots\ldots, b_n,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$k_0, k_1, k_2, \ldots\ldots\ldots\, k_n$$

called the "*coefficients,*" we are required to find out the numbers

$$x_1, x_2, \ldots\ldots\ldots, x_n, \ldots \qquad \ldots \qquad \ldots \quad (1)$$

satisfying the m conditions:

$$\left.\begin{array}{l} a_1x_1+a_2x_2+\ldots\ldots\ldots+a_nx_n=a_0 \\ b_1x_1+b_2x_2+\ldots\ldots\ldots+b_nx_n=b_0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ k_1x_1+k_2x_2+\ldots\ldots\ldots+k_nx_n=k_0 \end{array}\right\} \quad \ldots \qquad \ldots \quad (2)$$

The system (2) is called a *system of linear equations.* A linear equation is called *homogeneous,* if its term on the right side in (2) equals 0, and also the system (2) is called homogeneous, if every equation is homogeneous, *i.e.*, if $a_0=b_0 \ \ldots \ =k_0=0$. Every system of numbers (1) satisfying (2) is called a solution of (2).

At first we will try to get the solutions in the most simple cases.

I. $\quad n=m=1 \qquad\qquad a_1x=a_0$

(α) $a_1 \neq 0$. $\qquad$ Solution $x=a_0 : a_1$

(β) $a_1=0 \quad a_0 \neq 0$. $\qquad$ No solution.

(γ) $a_1=a_0=0$. $\qquad$ Solution x is arbitrary.

Remark: Division is permissible, if and only if the denominator is not equal to zero. If the denominator is not a constant, but a function, it is necessary to treat separately the cases in which this function vanishes. In the lecture, examples of this kind will be given.

II. $n=2,\ m=1$ $\qquad a_1x_1+a_2x_2=a_0$

(α) $a_1 \neq 0,\ a_2 \neq 0.$ Solutions (x_1, x_2); one of the numbers is arbitrary, the other is defined by it.

(β) $a_1 \neq 0,\ a_2=0.$ Solutions $(a_0:a_1,\ x_2)$ x_2 is arbitrary.

(β') $a_1=0,\ a_2 \neq 0.$,, $(x_1,\ a_0:a_2)$ x_1 ,,

(γ) $a_1=a_2=0,\ a_0 \neq 0.$ No solution.

(δ) $a_1=a_2=a_0=0.$ Solutions x_1 and x_2 are arbitrary.

III. $n=1,\ m=2$ $\qquad a_1x=a_0$

$$b_1x=b_0$$

(α) $a_1b_0-b_1a_0 \neq 0.$ No solution.

(β) $a_1b_0-b_1a_0=0,\ (a_1, b_1) \neq (0, 0).$

Solutions: $a_0 : a_1$, or $b_0 : b_1$, or $a_0 : a_1=b_0 : b_1$

if $a_1 \neq 0$ $\quad b_1 \neq 0$ $\quad a_1 \neq 0,\ b_1 \neq 0.$

(γ) $a_1=b_1=0.$ $(a_0, b_0) \neq (0, 0).$ No solution.

(δ) $a_1=b_1=a_0=b_0=0.$ Solution x is arbitrary.

IV. $n=2,\ m=2$ $\qquad a_1x_1+a_2x_2=a_0$

$$b_1x_1+b_2x_2=b_0$$

$$a_1b_2-a_2b_1=\Delta$$

$$a_0b_2-a_2b_0=\Delta_1$$

$$-a_0b_1+a_1b_0=\Delta_2$$

Necessary conditions for solutions (x_1, x_2) are $\Delta x_1 = \Delta_1$
$\Delta x_2 = \Delta_2$

(α) $\Delta \neq 0$. One solution $(\Delta_1 : \Delta, \Delta_2 : \Delta)$

(β) $\Delta = 0$, $(\Delta_1, \Delta_2) \neq (0, 0)$. No solution.

(γ) $\Delta = \Delta_1 = \Delta_2 = 0$, $a_1 b_2 (= b_1 a_2) \neq 0$.

Every solution of the first equation is also a solution of the second.—See II (α).

(δ) $a_1 = a_2 = 0$, $a_0 \neq 0$. No solution.

(δ') $b_1 = b_2 = 0$, $b_0 \neq 0$. No solution.

(ϵ) $a_1 = a_2 = a_0 = 0$. Every solution of the second equation is a solution of the system.—See II.

(ϵ') $b_1 = b_2 = b_0 = 0$. Every solution of the first equation is a solution of the system.—See II.

(ζ) $a_1 = b_1 = \Delta_1 = 0$. Solutions.—See III, β, γ, δ.

(ζ') $a_2 = b_2 = \Delta_2 = 0$. Solutions.—See III, β, γ, δ.

§ 2. The Homogeneous System belonging to an Arbitrary System of Linear Equations.

From the examples given in § 1 we may realize that even in the most simple cases a great number of different sub-cases has to be considered. The number of these sub-cases seems to increase infinitely with n and m. For avoiding these difficulties we will follow a way very characteristic of mathematical thoughts. We will suppose that solutions have been found out, and we will enquire about the connection between those solutions. By considering the properties of the system of all solutions we will get a general theory of the systems of linear equations including also different ways for finding out the solutions of a given system.

If $(x_1, x_2, \dots, x_n) = (w_1, w_2, \dots, w_n)$ and
$(x_1, x_2, \dots, x_n) = (u_1, u_2, \dots, u_n)$ are solutions of (2), then
$(x_1, x_2, \dots, x_n) = (w_1 - u_1, \dots, w_n - u_n)$ is a solution of

$$\begin{aligned} a_1 x_1 + \dots\dots + a_n x_n &= 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots & \\ k_1 x_1 + \dots\dots + k_n x_n &= 0 \end{aligned} \qquad (2/H)$$

and conversely, if $(u_1, \ldots, u_n)$ is a solution of (2), and $(y_1, \ldots, y_n)$ a solution of (2/H), then $(u_1 + y_1, \ldots, u_n + y_n)$ is a solution of (2). Therefore the following theorem holds:

Theorem I. Starting from an arbitrary solution of (2) we will get all solutions by addition of the solutions of (2/H).

The homogeneous system (2/H) *belongs* to the system (2). For solving (2) we have to find out all solutions of (2/H) and an arbitrary one of (2). For that purpose the introduction of a new notion is convenient.

§ 3.✓ THE *n-vectors*.

Definition 1: An ordered set of n numbers is called an n-vector.

$$\alpha = (a_1, \ldots\ldots, a_n) \tag{3}$$

The n numbers a_i defining the n-vector are called its co-ordinates. As this set is an ordered one, the vector will generally be changed by the interchange of the co-ordinates.

Examples: (1) The co-efficients of an arbitrary equation of (2/H) define an n-vector; it is called the "vector of that equation," and also the "vector of that row."

(2) The co-efficients of an arbitrary column of (2/H) define an m-vector, the "vector of that column."

(3) The solution (1) of (2) defines an n-vector, the "vector of the solution."

(4) Vectors in the plane (the space), in the sense this word is ordinarily used, are 2-vectors (3-vectors).

(5) Let n be the number of the customers of a bank; the balances of the customers are the co-ordinates of an n-vector representing the actual state of the bank. The reader may interpret the vector-addition defined below for this example.

Definition 2: The product of a number c and the vector α is an n-vector.

$$c\alpha = (ca_1, \ldots ca_n). \tag{4}$$

Definition 3: The sum of α and $\beta = (b_1, \ldots, b_n)$ is

$$\alpha + \beta = (a_1 + b_1, \ldots, a_n + b_n). \tag{5}$$

From these definitions it follows:

$$\begin{aligned} \alpha+\beta &= \beta+\alpha && \text{commutative law,} \\ \alpha+(\beta+\gamma) &= (\alpha+\beta)+\gamma && \text{associative law,} \\ c(\alpha+\beta) &= c\alpha+c\beta && \text{1}^{st} \text{ distributive law,} \\ (c_1+c_2)\alpha &= c_1\alpha+c_2\alpha && \text{2}^{d} \text{ distributive law.} \end{aligned} \tag{6}$$

As these laws hold, we can use the notations of sum of n-vectors in the same manner as it is to be used for numbers:

$$\Sigma\, c_i a^i = (\Sigma\, c_i a_1^i, \dots, \Sigma\, c_i a_n^i) \tag{7}$$

$i = 1, \dots, m$; c being arbitrary numbers; $a^i = (a_1^i, \dots, a_n^i)$

being arbitrary n-vectors.

The vector $-1 \cdot \alpha$ is called the *negative* of α and written $-\alpha$. The addition of $-\alpha$ is the inverse operation to the addition of α. As in elementary arithmetics this inverse operation is called *subtraction* and written by the sign $-$. Therefore:

$$\beta+(-\alpha) = \beta-\alpha\,. \tag{8}$$

The following special n-vectors will often be used:

$$\begin{aligned} 0 &= (0, \dots, 0) && \text{Zero-vector} \\ e^1 &= (1, 0, \dots, 0) && \text{1}^{st} \text{ Unit-vector} \\ e^2 &= (0, 1, 0, \dots, 0) && \text{2}^{d} \text{ Unit-vector} \\ &\dots\dots\dots\dots \\ e^n &= (0, 0, \dots 0, 1) && n^{th} \text{ Unit-vector.} \end{aligned} \tag{9}$$

Formulae:

$$\begin{aligned} \alpha-\alpha &= 0 \\ c \cdot 0 &= 0 \\ \alpha &= \Sigma\, a_i e^i \end{aligned} \tag{10}$$

§ 4. Vector-spaces.

Definition 3: The n-vector $\Sigma c_i \alpha^i$ is *dependent* on the n-vectors α^i.

Definition 4: The n-vectors $\alpha^1, \ldots\ldots, \alpha^m$ are said to be *independent* if none of them is dependent on the $m-1$ others and $m > 1$. A single n-vector is independent, if it is $\neq 0$.

Definition 5: The set of all vectors dependent on the α^i is called the *vector-space* generated by the α^i.

Definition 6: A set of independent n-vectors generating a vector-space is called its *Basis*.

Theorems concerning vector-spaces:

1. The n-vectors α^i are independent if and only if for every system of numbers $(c_1, \ldots, c_m) \neq (0, \ldots, 0)$ we have $\Sigma c_i \alpha_i \neq 0$.

Proof. 1. If $c_i \neq 0$, α^i is dependent on the other α^j. 2. If α^j is dependent on the other α^i, then $\alpha^j = \Sigma d_p \alpha^p$, and therefore $\Sigma c_i \alpha^i = 0$, for $c_j = -1$, $c_p = d_p$. The theorem is evident in the case of a single n-vector.

but they do not form the basis

2. If $\beta^1, \ldots, \beta^r$ belong to a vector-space V, every n-vector κ dependent on the β^i belongs also to V.

Proof. Let V be generated by $\alpha^1 \ldots \alpha^m$, then from $\kappa = \Sigma k_j \beta^j$, $\beta^j = \Sigma b^j_i \alpha^i$ follows $\kappa = \Sigma c_i \alpha^i$, where $c_i = \Sigma k_j b^j_i$.

3. Every vector-space containing n-vectors $\neq 0$ has a basis.

Proof. If the vectors α^i generating V are not independent, α^j may depend on the other $m-1$ generating n-vectors. From 2 it follows that these vectors generate also V. On repeating—if it is necessary—this reduction we will get after a finite number of steps a subset of the α^j generating V composed of independent vectors, *i.e.*, a basis of V.

4. If $\alpha^1, \ldots, \alpha^m$ is a basis of V, $\beta = \Sigma c_i \alpha^i$, and $c_j \neq 0$, then we will get a new basis of V on replacing α^j by β.

Proof. The vector-space generated by β and the $\alpha^{i \neq j}$ is contained in V. On the other hand it contains $\beta - \sum_{k \neq j} c_k \alpha^k = c_j \alpha^j$ and therefore α^j. We have to show that the m n-vectors β and $\alpha^{i \neq j}$ are independent.

Let $d\beta + \sum_{k \neq j} d_k \alpha^k = 0$, on replacing β by its value $\Sigma c_i \alpha^i$ we get a vanishing linear function of the α^i whose coefficients vanish, as the α^k are independent. The coefficient of α^j is dc_j; as $c_j \neq 0$, it follows: $d=0$. As the α^k are independent, the d_k are vanishing. Therefore the α_k and β are independent; so they form a basis of V.

5. If $\alpha^1, \ldots, \alpha^m$ is a basis of V, and $\beta^1, \ldots, \beta^t$ are independent vectors in V, then we get a new basis on replacing t suitable α^i by the β, and $t \leq m$ holds.

Proof. From 4 it follows that a suitable α can be replaced by β^1. Let $\alpha^1, \ldots, \alpha^{t}, \ldots, \beta^1 \ldots, \beta^r$ be a basis of V, $r<t$ then we can express β^{r+1} by these m vectors, and in this expression the coefficient of at least one α do not vanish; therefore this α can be replaced by β^{r+1}. Therefore we can continue replacing an α by a β till $r=t$, or $r=m$. In the last case all β's must depend on the $\beta^1, \ldots, \beta^m$, and therefore $t=m$ in this case. Generally $t \leq m$.

Definition 7: The maximum number of independent vectors of V is called the *Rank* of V.

6. The number of the vectors of an arbitrary basis of V equals the rank of V. (Therefore every basis of V has the same number of vectors.)

Proof. From 5 it follows that there cannot be more independent n-vectors in V than an arbitrary basis has elements.

7. If every n-vector of V belongs to V′, but not every n-vector of V′ belongs to V, then is the rank of V less than the rank of V′.

Proof. Let $\alpha^1, \ldots, \alpha^r$ be a basis of V, and β a vector of V′ not contained in V, then is r the rank of V, but the rank of V′ is at least $r+1$ because V′ contains $r+1$ independent elements α^i, β.

8. The rank of a vector-space of n-vectors is at most n.

Proof. The n unit-vectors (9) generate a vector-space V′ in which the co-ordinates are arbitrary numbers; therefore V′ contains every n-vector, and 8 follows from 7.

9. Between $p>n$ of n-vectors there exists always a linear equation with non-vanishing coefficients.

Proof. If these n-vectors are independent they will generate an n-vector-space of rank $p>n$.

10. A system A of n-vectors, with the property that the sum of two arbitrary elements of A, and also the product of an arbitrary element of A with an arbitrary real number belongs to A, is a vector-space.

Proof. From (9) it follows that in A there exists a maximum number $0 \leq r \leq n$ of independent n-vectors. The vector-space generated by such $r>0$ independent n-vectors is identical with A. If $r=0$, A contains only the n-vector 0.

§ 5. THE VECTOR-SPACES CONNECTED WITH A SYSTEM OF HOMOGENEOUS LINEAR EQUATIONS.

Theorem II. The solutions of (2/H) form a vector-space X.

Proof. From $\Sigma a_i x_i = \Sigma b_i x_i = \ldots = \Sigma k_i x_i = 0$

and $\Sigma a_i y_i = \Sigma b_i y_i = \ldots = \Sigma k_i y_i = 0$ it follows

$$\Sigma a_i c x_i = \Sigma b_i c x_i = \ldots = \Sigma k_i c x_i = 0 \quad \text{and}$$

$$\Sigma a_i (x_i + y_i) = \Sigma b_i (x_i + y_i) = \ldots = \Sigma k_i (x_i + y_i) = 0$$

Therefore the solutions satisfy the conditions of a vector-space given in §4, 10.

Theorem III. To every vector-space X there exists a vector-space V, such that every n-vector of X is a solution of a linear homogeneous equation, if and only if the vector of this equation is a vector of V.

Proof. The vectors of X are solutions of the equation defined by the n-vector 0. If they are solutions of the equations defined by α and by β, then they are also solutions of the equations defined by $c\alpha$ and by $\alpha+\beta$. From §4, 10 it follows that the set of the vectors α, β, ... form a vector-space.

The Theorems II and III show that our problem (2/H) is closely connected with two vector-spaces X and V. The vectors of (2/H) generate a vector-space V′ and every vector of V′ is also a vector of V. If V and V′ were not identical, then V should be of higher rank. In that case two vector-spaces of equation-vectors of different rank would define the same

vector-space X. In other words there would exist an n-vector λ independent of the vectors of the equations (2/H), such that every solution of (2/H) is also a solution of the equation generated by λ. We will see later that this case is impossible.

§ 6. THE BASIS OF A VECTOR-SPACE. THE METHOD OF "SWEEP-OUT."

For further consideration of linear equations we need know some more properties of vector-spaces.

11. If $\alpha, \beta, \dots, \kappa$ are n-vectors generating a vector-space V, $a \neq 0$ and b are arbitrary numbers, then $a\alpha + b\beta, \beta, \dots, \kappa$ generate also V.

Proof. As V includes $a\alpha + b\beta$, and the vector-space generated by $a\alpha + b\beta$, $\beta, \dots, \kappa$ includes α, the two vector-spaces are identical.

12. If the n-vectors $\alpha, \beta, \dots, \kappa$ generating V are not all equal to 0, then we will get another system of n-vectors generating V by omitting the n-vectors equal to 0.

Proof. The n-vector 0 is dependent on the other ones.

13. The n-vectors

$$\begin{aligned} \beta^1 &= (-1, 0, \dots, 0, \ \ 0, b^1_{r+1}, \dots, b^1_n) \\ \beta^2 &= (0, -1, \dots, 0, \ \ 0, b^2_{r+1}, \dots, b^2_n) \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \beta^r &= (0, \ \ 0, \dots, 0, -1, b^r_{r+1}, \dots, b^r_n) \end{aligned} \quad \dots \quad (11)$$

are independent.

Proof. Let $\Sigma c_i \beta^i = 0$, the co-ordinates of the vector on the left are equal to the co-ordinates on the right hand. Therefore $c_1 = c_2 = \dots = c_r = 0$.

14. V may be an arbitrary vector-space including n-vectors $\neq 0$; starting from an arbitrary finite system of n-vectors generating V and using the methods of 11 and 12 we will get after a finite number of suitable steps a basis of V differing from the n-vectors of formula (11) at most by a permutation of the co-ordinates of all vectors.

Remark: The methods 11 and 12 can be considered as operations practised on the scheme of co-ordinates written on the right hand in formula

(11). It is therefore useful to describe 14 as a property of that scheme. For this purpose we introduce a notation often used in mathematics.

Definition 8: The system of the co-ordinates of m ordered n-vectors is called a *Matrix* M. The co-ordinates of the s^{th} vector form the s^{th} *row* of M; the t^{th} co-ordinates of the n-vectors form the t^{th} *column* of M $(1 \leq s \leq m, 1 \leq t \leq n)$. M$=0$ means that every co-ordinate of M vanishes.

The operation given by 11 and 12 may be shortly expressed by the words "row-addition" and "omission of 0-rows." Using these words we get a theorem (14′) equivalent to (14).

14′. By the operations of row-addition and omission of 0-rows every matrix M $\neq 0$ can be transformed to a matrix of the kind (11), or to a matrix differing from it only by a permutation of the columns. The proof of 14′—and therefore also the proof of 14—will be made by different steps. To simplify the description of these transformations it may be understood that every 0-row that would appear should be omitted automatically, without mentioning that operation. Every row contains therefore at least one co-ordinate $\neq 0$. For abbreviation we will call

the rows of the matrix : (1), (2),.........

the co-ordinates of (i) : $[1, i], [2, i], \ldots, [n, i]$; $i = 1, 2.$ ... (12)

the columns : $\langle 1\rangle, \langle 2\rangle, \ldots, \langle n\rangle$.

These signs do not denote constant values, they change at every step of our transformation, *e.g.*, (1) denotes the first row at every step, also $\langle i_m\rangle$ the i_m^{th} column, i_m being an arbitrary number, etc.

On using these signs the different steps of the transformation may be described as follows :

1, *a*) i_1 may be the smallest number such that $[i_1, 1] \neq 0$,

by the row-addition $(1) \to \{-1 : [i_1, 1]\}\ (1)$

$[i_1, 1]$ becomes -1 ;

1, *b*) by the row-addition $(k) \to (k) + [i_1, k]\ (1)$

$[i_1, k]$ becomes 0

By these operations the column $< i_1 >$ has been "swept out," *i.e.*, one of the numbers has been made -1, all others vanish. We continue on sweeping out:

2, *a*) i_2 may be the smallest number such that $[i_2, 2] \neq 0$, then is $i_2 \neq i_1$, by row addition $(2) \longrightarrow \{-1 : [i_2, 2]\}$ (2)

$[i_2, 2]$ becomes -1, and $[i_1, 2] = 0$ is not changed.

2, *b*) By the row-addition $(k) \longrightarrow (k) + [i_2, k]$ (2)

for every $k \neq 2$, $[i_2, k]$ becomes 0, and $<i_1>$ is not changed.

Now the columns $<i_1>$ and $<i_2>$ are "swept out." We continue this sequence of operations. After $2(q-1)$ steps the columns

$<i_1>, <i_2>, \ldots, <i_{q-1}>$ may be "swept out," *i. e.*

$[i_s, s] = -1$, $[i_s, k] = 0$, for $s = 1, \ldots q-1$, $k \neq s$, $\quad i_1, \ldots, i_{q-1}$

being different numbers. If after these steps (and the automatic omitting of 0-rows) the matrix has more than $q-1$ rows,

q, *a*) i_q may be the smallest number for which $[i_q, q] \neq 0$;

by $(q) \longrightarrow \{-1 : [i_q, q]\}$ (q) $\quad [i_q, q]$ becomes -1

q, *b*) By $(k) \longrightarrow (k) + [i_q\, q]$, $k \neq q$, $[i_q, k]$ becomes 0;

the rows $<i_1>, \ldots, <i_{q-1}>$ are not changed by these transformations.

By mathematical induction it follows that the method can be continued till the number of columns swept out equals the number of the remaining rows. After this transformation

$<i_1>, \ldots, <i_r>$ may be swept out, *i. e.* $[i_s, s] = -1$, $[i_s, t] = 0$,

for $s = 1, \ldots, r$, $t \neq s$. By a permutation of the columns, transforming $i_s \longrightarrow s$, the matrix will take the form (11).

Corollary: Starting from an arbitrary system of n-vectors generating $V \neq 0$ it is always possible to get a basis of V.

Proof: Let m be the number of the generating vectors, it is always possible to sweep out the matrix by at most m^2 row-additions and omissions of at most $m-1$ rows. As the n-vectors of the remaining rows are independent, they form a basis.

§ 7. Solution of Systems of Homogeneous Linear Equations by the Method of "Sweep-out."

Theorem IV. Let V′ be the vector-space generated by the n-vectors of (2/H), X the vector-space of the solutions of (2/H), and V the vector-space defined by Theorem III, (13)
then V = V′, rank (V) + rank (X) = n holds.

Theorem V. It is possible to find out a basis of X by a limited number of row-additions and omissions of 0-rows.

Proof of the Theorems IV and V: Every solution of (2/H) is a solution of the equations belonging to the n-vectors of V′ and *vice versa*. To obtain these solutions it is sufficient to find out the solutions of the equations belonging to a basis of V′. As it was proved by 14 it is always possible to find out such a basis by the method of sweep-out. This basis can be taken in the form (11), only by a permutation $i_k \longrightarrow k$ of the columns. If we write therefore $x_{i_k} = y_k$, $k=1,\dots\dots, n$

then our problem will be reduced to the following :

$$\begin{array}{c} -y_1 + b^1_{r+1} y_{r+1} + \dots + b^1_n y_n = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ -y_r + b^r_{r+1} y_{r+1} + \dots + b^r_n y_n = 0 \end{array} \qquad (14)$$

or

$$y_s = \sum_{t=1}^{n-r} b^s_{r+t} y_{r+t} \qquad s=1,\dots,r.$$

The $y_{r+1},\dots, y_n$ can take arbitrary independent values, and $y_1,\dots, y_r$ are uniquely determined by them. There exists therefore a (1, 1) correspondence between the vector-space of all $(n-r)$-vectors $(y_{r+1},\dots,y_n)$, and the vector-space of all solutions $(y_1,\dots, y_n)$

$$\begin{array}{ccc} (1, 0,\dots, 0, 0) & & (b^1_{r+1},\dots, b^r_{r+1}, 1, 0,\dots 0, 0) \\ \dots\dots\dots\dots & \text{and} & \dots\dots\dots\dots\dots\dots\dots\dots \\ (0, 0,\dots, 0, 1) & & (b^1_n ,\dots, b^r_n , 0, 0,\dots, 0, 1) \end{array}$$

are corresponding vectors. As this correspondence is invariant for vector-addition and multiplication with a number, and the $(n-r)$-vectors on the right form a basis of the solutions (14), we will get a basis of X by the permutation $s \longrightarrow i_s$ of the co-ordinates.

Therefore Theorem V is true, and

$$\text{Rank } (V') + \text{rank } (X) = n$$

holds. As X is also the vector-space of all solutions of V we can replace V′ by V in this formula ; hence (13) holds. Every n-vector of V′ belongs to V and these vector-spaces have the same rank ; therefore (§4, 2) V and V′ are identical.

§ 8. Solutions of Non-homogeneous Linear Systems.

In order to solve (2) it would be good to consider (2), (2/H), and

$$\begin{array}{c} a_1x_1 + \dots\dots + a_nx_n + a_0x_0 = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ k_1x_1 + \dots\dots + k_nx_n + k_0x_0 = 0 \end{array} \qquad (15)$$

the generating $(n+1)$-vectors being called :

$\bar{\alpha} = (a_1, \dots, a_n, a_0) \dots\dots, \quad \bar{\kappa} = (k_1, \dots k_n, k_0)$.

Theorem VI : The system (2) is solvable if and only if the vector-space V generated by the n-vectors $\alpha, \dots, \kappa$ and the vector-space $\bar{\text{V}}$ generated by the $(n+1)$-vectors $\bar{\alpha}, \dots \bar{\kappa}$ have the same rank. If (2) is not solvable, then rank $\bar{\text{V}} = 1 + \text{rank V}$.

Proof : As every linear homogeneous relation between the $\bar{\alpha} \dots \bar{\kappa}$ holds, when $a_0, \dots, k_0$ will be replaced by 0, the rank r of V is not greater than the rank $\bar{r}$ of $\bar{\text{V}}$. Let X and $\bar{\text{X}}$ be the vector-spaces generated by the solutions of (2) and (15), then rank $\text{X} = n - r$, rank $\bar{\text{X}} = n + 1 - \bar{r}$. If $(x_1, \dots, x_n)$ is a solution of (2/H), then $(x_1, \dots, x_n, 0)$ is a solution of (15). Therefore rank $\text{X} \leqq$ rank $\bar{\text{X}}$, $r + 1 \geqq \bar{r} \geqq r$. If $\bar{r} = r + 1$, rank $\bar{\text{X}} =$ rank X, then (15) has only solutions with $x_0 = 0$, and (2) has no solutions. If $\bar{r} = r$, rank $\bar{\text{X}} >$ rank X, and there exists a solution $(x_1, \dots x_n, x_0 \neq 0)$ of (15). Hence $\left(\frac{-x_1}{x_0}, \dots, \frac{-x_n}{x_0}\right)$ is a solution of (2).

Theorem VII : If there are solutions of (2), we will get them by "sweep-out."

Proof : We can get a basis of the solutions $\bar{\text{X}}$ of (15) by the method of sweep-out. If there is a solution in which $x_0 \neq 0$, it must also be a basis $(n+1)$-vector of this kind. From this $(n+1)$-vector we will get a solution of (2) on dividing the co-ordinates by $-x_0$.

As *0* is a solution of (2) if and only if the equations are homogeneous, the solutions form a vector-space only in this case. In general the set of all solutions will be described by the following theorem :

Theorem VIII: If ξ and η are solutions of (2), and $s+t=1$, then $s\xi+t\eta$ is also a solution. *Vice versa* if a set W of n-vectors has the property that with two n-vectors ξ and η the n-vector $s\xi+t\eta$ belongs to W, then W is the set of the solutions of a system (2).

Proof: From $\Sigma a_i x_i = a_0$, $\Sigma a_i y_i = a_0$ it follows that $\Sigma a_i\ (sx_i+ty_i) = (s+t)\ a_0 = a_0$. As the same holds for the other equations of (2), the first proposition is proved.—If χ, λ, π are n-vectors of W, then $2\ (1/2\ \chi + 1/2\ \lambda) - \pi = \mu$, and $c\chi + (1-c)\ \pi = \nu$ belongs to W. Hence $(\chi-\pi)+(\lambda-\pi)=\mu-\pi$, $c\ (\chi-\pi)=\nu-\pi$. If π is a fixed element of W, the set of all differences $(\chi-\pi)$, $(\lambda-\pi)$ is a vector-space X (§ 4, 10). From Theorem III it follows that X is the set of the solutions of a system of linear homogeneous equations. This system may be (2/H). If $\pi=(p_1, \ldots, p_n)$, and we define $a_0, \ldots, k_0$ by $a_0=\Sigma p_i a_i, \ldots, k_0=\Sigma p_i k_i$, then, by Theorem I, W is the set of all solutions of (2).

§ 9. The Method of Orthogonalization.

By the previous theorems the problem proposed in the introduction has been solved completely. It will always be the principal part of the calculation, to find out a basis of the vector-space X. We may do it by the method of sweep out, but it is important to have other suitable methods for it. In this section the method of orthogonalization will be treated ; its advantage is that we will get at the same time a basis of X and a basis of V, both in a special form.

In the previous, and in the following paragraphs there is no restriction about the numbers, which should be used.[1] However in this section we will suppose that all numbers used are real.

Definition 9: The *Scalar product* $S.\alpha\beta$ of two n-vectors $\alpha=(a_1, \ldots, a_n)$ and $\beta=(b_1, \ldots, b_n)$ is:

$$S.\alpha\beta = S.\beta\alpha = \Sigma a_i b_i. \qquad \ldots \quad (16)$$

Definition 10: The *Length* of α is a number $|\alpha| \geqq 0$, defined by

$$|\alpha|^2 = S.\alpha\alpha \qquad \ldots \quad (17)$$

[1] On using the notation that will be introduced in Chapter II : The theorems and the proofs hold for an arbitrary field of characteristic 0.

Formulae: $|\alpha| = 0$, if and only if $\alpha = 0$.

$$S.\alpha(\beta+\gamma) = S.\alpha\beta + S.\alpha\gamma$$

$$S.c\alpha\beta \quad = cS.\alpha\beta \qquad \dots \quad (18)$$

$$-|\alpha|\,|\beta| \leq S.\alpha\beta \leq |\alpha|\,|\beta| \quad \text{(Cauchy's inequality)}$$

$$|\alpha+\beta| \leq |\alpha| + |\beta|$$

Definition 11: If $S.\alpha\beta = 0$, α and β are *orthogonal*.

Remark: In the lectures the geometrical significance of these notions and formulae will be explained.

The co-ordinates of α can be expressed as scalar products:

$$a_i = S.\alpha\epsilon^i \qquad \dots \quad (19)$$

Definition 12: The vectors $\beta^1, \dots, \beta^s$ form an *Orthogonal System* if they satisfy the conditions:

$$S.\beta^i\beta^i = 1$$

$$S.\beta^i\beta^j = 0 \qquad i \neq j. \qquad \dots \quad (20)$$

Properties of orthogonal systems:

1. If the β^i form an orthogonal system, they are independent.

Proof: From $0 = \Sigma c_i\beta^i$ it follows

$$0 = \beta^k\Sigma c_i\beta^i = c_k \text{ for } k = 1, \dots, s.$$

2. If the β^i form an orthogonal system, and χ is independent of the β^i, then there is an n-vector β^{s+1}, such that χ is dependent on $\beta^1, \dots, \beta^{s+1}$, and these n-vectors form an orthogonal system.

Proof: $\lambda = \chi - \Sigma(S.\beta^i\chi)\,\beta^i$, and $\beta^{s+1} = |\lambda|^{-1}\lambda$, then there is for $k = 1, \dots, s$, $0 = S.\beta^k\lambda = S.\beta^k\beta^{s+1}$, $S.\beta^{s+1}\beta^{s+1} = 1$.

3. If V is a vector-space of rank m, containing a vector-space A of rank $r < m$, then there exists a basis of V, forming an orthogonal system $\beta^1, \dots, \beta^r, \dots, \beta^m$, such that $\beta^1, \dots \beta^r$ form a basis of A.

Proof: Let β^1 be an arbitrary n-vector of A of the length 1; if $r > 1$ there is in A an n-vector independent of β^1, and therefore there is an orthogonal system β^1, β^2. By repeated use of this construction, we get $\beta^1, \dots, \beta^r$ in A and by continuing this method in V we will get $\beta^{r+1}, \dots, \beta^m$.

Theorem IX: Let V be the vector-space generated by the vectors of the equations (2/H), and X the vector-space of the solutions of (2/H), then there exists an orthogonal system $a^1, \ldots, a^r, \xi^1, \ldots, \xi^{n-r}$, such that $a^1, \ldots, a^r$ form a basis of V, and $\xi^1, \ldots, \xi^{n-r}$ form a basis of X.

Proof: The connection between the vector-spaces V and X is that every n-vector of X is orthogonal to every n-vector of V (and *vice versa*). Hence if we construct an orthogonal system of n n-vectors; $a^1, \ldots, a^r, \xi^1, \ldots, \xi^{n-r}$, such that the first r n-vectors form a basis of V, the $\xi^1, \ldots, \xi^{n-r}$ are orthogonal to the n-vectors of V, and form therefore solutions of (2/H). An arbitrary n-vector can be expressed by $\lambda = \Sigma c_i a^i + \Sigma d_j \xi^j$. λ is orthogonal to the n-vectors of V, if and only if for $k=1, \ldots, r$ $0 = \mathbf{S}.\lambda a_k = c_k$ holds. $\xi^1, \ldots, \xi^{n-r}$ is therefore a basis of X.

Remark: The proof of Theorem IX does not apply to the Theorems IV-VIII and the method of sweep-out; Theorem IV follows directly from Theorem IX.

§ 10. Substitution and Elimination.

The method of Substitution. In order to get the solutions of (2) we can also reduce the problem to $m-1$ equations and $n-1$ unknown. Let x_{i_1} be the first unknown for which the coefficient of the 1st equation does not vanish, then:

$$x_{i_1} = \sum_{i \neq i_1} \frac{-a_i}{a_{i_1}} x_i + \frac{a_0}{a_{i_1}}$$

On substituting this value in the other $m-1$ equations we get $m-1$ equations with $n-1$ unknown. If some of these equations become identical they will be omitted. The rows of the new equations are

$$(k) - [i_1, k] : a_i (1);$$

hence the substitution of x_{i_1} is nothing more than the sweep-out of the column $\langle i_1 \rangle$. On continuing this procedure the column $\langle i_2 \rangle$ will be swept out in the rows (2), ..., (m), also $\langle i_3 \rangle$ in the rows (3), ...(m); etc. the 0-rows, as they belong to identical equations, being always omitted. The result will be that the columns $\langle i_1 \rangle, \ldots, \langle i_r \rangle$ will be swept out in the rows below $[i_s, s]$. The last of these equations give us the possibility to express x_{i_r} by the remaining $n-r$ unknown. The substitution of this

value in the $r-1$ equations is nothing but the sweep-out of the column $<i_r>$ in the rows above the row (r). On repeating this procedure by the substitution of $x_{i_{r-1}}, \ldots, x_{i_2}$ expressed by the last $n-r$ unknown we sweep out the corresponding columns in the upper rows. Hence the method of substitution is not quite different from the method of sweep-out.

The problem of Elimination. The equations (2/H) will be solved, when we have got equations of the kind (14). As the n-vectors belonging to (14) are dependent on the vectors belonging to (2/H), we should multiply the equations (2/H) with suitable numbers and add so that the sum of these equations becomes an equation of the kind (14), *i.e.*, we *eliminate* the unknown $x_{i_2}, \ldots, x_{i_r}$ to get the first row of (14), $x_{i_1}, x_{i_3}, \ldots, x_{i_r}$ to get the second row, etc. By this procedure we intend to find out at once what is done step by step on using the method of sweep-out. The rank of V and also the factors we need for the elimination, are dependent only on the given coefficients. We have to consider the character of that dependence. For this purpose we need the "determinants."

§ 11. PRINCIPAL PROPERTIES OF THE DETERMINANTS.

Definition 13: A function

$$L = L(a^1, \ldots, a^n) \qquad \ldots \quad (21)$$

of n n-vectors

$$a^i = (a^i_1, \ldots, a^i_n)$$

is called a *Determinant*

$$\begin{vmatrix} a^1_1 \ldots\ldots a^1_n \\ \ldots\ldots\ldots\ldots \\ a^n_1 \ldots\ldots a^n_n \end{vmatrix} = \det a^i_k \qquad \ldots \quad (22)$$

if the following conditions are satisfied:

(*a*) $L(a^1, \ldots, ca^m, \ldots, a^n) = c\, L(a^1, \ldots, a^n)$, for $m = 1, \ldots, n$,

(*b*) L will not change by replacing

$$a_i \longrightarrow a_k + a_i, \text{ for } i \neq k, \qquad \ldots \quad (23)$$

(*c*) $L(e^1, \ldots, e^n) = 1$.

It will be proved later on that a function of this kind exists, and that it is uniquely defined by the conditions (*a*), (*b*), (*c*). Now we will consider

the properties of a function—if there is one—satisfying the above conditions.

1. If $a^i=0$, $L=0$.

Proof: $0=0 \cdot L=L(a^1, \ldots, 0 \cdot a^i, \ldots, a^n)=L(a^1, \ldots, a^n)$.

2. L will not be changed by replacing $a^k \longrightarrow a^k + ca^i$, if $a^i \neq a^k$, and c is an arbitrary number.

Proof:
$$L = \frac{1}{c} L(a^1, \ldots, ca^i, \ldots, a^k, \ldots a^n)$$
$$= \frac{1}{c} L(a^1, \ldots, ca^i, \ldots, a^k+ca^i, a^n)$$
$$= L(a^1, \ldots, a^i, \ldots, a^k+ca^i, \ldots, a^n).$$

3. On interchanging a^i and a^k L will be replaced by $-L$.

Proof: $L = L(a^1, \ldots, a^i, \ldots, a^k+a^i, \ldots, a^n)$
$$= L(a^1, \ldots, a^i-(a^k+a^i), \ldots, a^k+a^i, \ldots, a^n)$$
$$= L(a^1, \ldots, -a^k, \ldots, -a^k+a^k+a^i, \ldots, a^n)$$
$$= L(a^1, \ldots, -a^k, \ldots, a^i, \ldots, a^n)$$
$$= -L(a^1, \ldots, a^k, \ldots, a^i, \ldots, a^n).$$

4. If a^i is dependent on the $n-1$ other n-vectors, $L=0$.

Proof: From (2) it follows that a^i can be replaced by 0. Hence $L=0$.

5. $L(e^{i_1}, \ldots, e^{i_n}) = 1$, if $i_1, \ldots, i_n$ is an even permutation,
$= -1$, if „ „ „ is an odd permutation,
$= 0$, if the indices are not different.

These formulae follow directly from (c), (3), and (4).

6. $L(a^1, \ldots, a^i+\beta, a^{i+1}, \ldots, a^n) = L(a^1, \ldots, \beta, \ldots, a^1)+L$.

Proof: Either a^i is dependent on the other n-vectors a^k, or these $n-1$ n-vectors are not independent, or the n-vectors a are all independent. In the first case $L=0$, on the left side a^i may be omitted, and therefore the equation holds. In the 2^d case each of the three functions equals zero. In the 3^d case β is dependent on the a^i. $\beta=\Sigma b_i a^i$. By applying 2 on both sides of the equation, the left side becomes $(1+b_i)L$, and the right side becomes $L+b_i L$; therefore 6 holds.

7. If $a^i = \sum_{s=1}^{t} c_s \beta^s$, and B_s is the determinant, we get by replacing $a^i \longrightarrow \beta^s$, i being constant, and $s=1, \dots, t$, then $L = \Sigma c_s B^s$.

Proof: To prove this theorem we have to use 6 and (a) t times.

8. Let L^i_k be the determinant we get by replacing a^i by e^k, and

let $a^i = \sum_k a^i_k e^k$, then

$$L = \Sigma a^i_k L^i_k. \quad \dots \quad \dots \quad (24)$$

The theorem follows directly from (7).

$$9. \; L = \Sigma \pm a^1_{k_1} \dots a^n_{k_n} \quad \dots \quad \dots \quad \dots \quad \dots \quad (25)$$

the summation has to be made only for the permutations $k_1, \dots, k_n$ of $1 \dots n$, + being valid for even and − for odd permutations.

Proof: On using 8 n times we get $L = \Sigma a^1_{k_1} a^2_{k_2} \dots a^n_{k_n} L(e^{k_1}, \dots, e^{k_n})$; hence from 5 follows the theorem.

Theorem X. The determinant $L(a^1, \dots, a^n)$ exists and is uniquely defined by every ordered set of n n-vectors $a^1, \dots, a^n$.

Proof: As it was stated in 9 the value of the determinant—if there exists one—cannot be different from (25). Therefore we have to prove that (25) has the properties (a), (b), (c). As (a) and (c) are obviously satisfied, we will prove (b). The right side of (25) is a function of $a^1, \dots, a^n$; it may be called $D(a^1, \dots, a^n)$. By interchanging two of the n-vectors of the argument the even permutations become odd and *vice versa*; therefore D is changed to −D. Hence if $a^i = a^k$, for $i \neq k$, then $D = -D = 0$. From (25) it follows that $D(a^1, \dots, (a^i + a^k), \dots, a^k \dots, a^n) = D(a^1, \dots, a^i, \dots, a^k, \dots a^n) + D(a^1, \dots, a^k, \dots, a^k, \dots, a^n)$, and as the last value vanishes, (b) is satisfied, and the theorem holds.

§ 12. Further Properties of the Determinants.

The columns of (22) form n-vectors:

$$a_1 = (a^1_1, a^2_1, \dots, a^n_1) \quad \dots \quad (26)$$

$$\dots\dots\dots\dots\dots\dots$$

$$a_n = (a^1_n, a^2_n, \dots, a^n_n).$$

10. $L(a^1, \dots, a^n) = L(a_1, \dots, a_n)$. ... (27)

i.e., a determinant will not be changed, if the rows are interchanged by the columns.

Proof: Let $k_1, \dots, k_n$ be the permutation inverse to $i_1, \dots, i_n$, then these permutations are either both even or both odd ;

$$a^1_{i_1}, \dots, a^n_{i_n} = a_1^{k_1}, \dots, a_n^{k_n}, \text{ and}$$

∴ $i \cdots$ is inverse permutation to $k_1 \cdots k_n$.

$$L(a^1, \dots, a^n) = \Sigma \pm a^1_{i_1} \cdot \dots, a^n_{i_n} = \Sigma \pm a_1^{k_1} \dots, a_n^{k_n} = L(a_1, \dots, a_n).$$

a', b', 1' to 9.') The formulae : (a), (b), (1) to (9) hold if we replace the a^i by the a_i.

These theorems are direct consequences of (11).

11. If we replace a^p by e^q, the determinant becomes independent of the co-ordinates of a_q.

Proof: On replacing a^p by e^q the product $a^1_{i_1}, \dots, a^p_{i_p}, \dots, a^n_{i_n}$ becomes 0, for $i_p \neq q$, hence in every non-vanishing product there is no a^s_q for $s \neq p$; a^p_q is replaced by 1, therefore the determinant becomes independent of the co-ordinates a^s_q of a_q.

The determinant we get on replacing $a^{p_1}, \dots, a^{p_m}$ by $e^{q_1}, \dots, e^{q_m}$, $p_1, \dots, p_m$ being different, and $q_1, \dots, q_m$ being different, will be called

$$L^{p_1, \dots, p_m}_{q_1, \dots, q_m} \quad \dots (28)$$

12. Let L' be the determinant we get on replacing in $L(a_1, \dots, a_n)$ the arguments $a_{q_1}, \dots, a_{q_m}$ by $e^{p_1}, \dots, e^{p_m}$, then

$$L' = L^{p_1, \dots, p_m}_{q_1, \dots, q_m} \quad \text{holds.}$$

Proof: From 10, it follows that (28) is not changed if we replace the $a^s_{q_i}$ ($s \neq p_i$) by 0. From 10 and 11 it follows also that L' is not changed if we replace the $a^{p_i}_s$ ($s \neq q_i$) by 0. In both cases we get a determinant with

the co-ordinates:

$$[k, i] = a_k^i \quad \text{for } \begin{matrix} i \neq p_1, \dots, p_m \\ k \neq q_1, \dots q_m \end{matrix}$$

$$= 1 \quad \text{for } (i, k) = (p_s, q_s), \; s = 1, \dots, m \qquad \dots \quad (29)$$

$$= 0 \quad \text{for } i = p_s, \; k \neq q_s$$

$$\text{or } i \neq p_s, \; k = q_s.$$

13. By a permutation of the $p_1, \dots p_m$ and a permutation of the $q_1, \dots q_m$ (28) is not changed, when these permutations are either both even, or both odd; (28) is changed to its negative value, when the two permutations are of different kind.

Proof: To every permutation of p (of q) belongs a corresponding permutation of the rows (of the columns) of (29). Hence the theorem holds.

If $n > 2$ we can select, to every combination of m positive integers $\leq n$, an even permutation

$$s_1, \dots, s_m, t_1 \dots, t_{n-m} \qquad \dots \quad (30)$$

so that s_i form the given combination. This selection being fixed for every of the $\binom{n}{m}$ combinations the following theorem holds:

14. $$\Sigma \, L^{p_1, \dots, p_m}_{s_1, \dots, s_m} \cdot L^{r_1, \dots, r_{n-m}}_{t_1, \dots, t_{n-m}} \qquad \dots \quad (31)$$

$= L$, when $p_1, \dots, p_m, r_1, \dots, r_{n-m}$ is an even permutation of $1, \dots n$

$= -L$, ,, ,, ,, ,, ,, ,, odd permutation of $1, \dots n$

$= 0$,, ,, ,, ,, ,, ,, are not n different numbers.

Proof: Let $p_1, \dots, p_m, r_1, \dots, r_{n-m}$ be an even permutation of $1, \dots, n$. The factors of each term of (31) may be developed as functions of the original a_k^i according to (25). The first factor is composed of $m!$ monomials of degree m, the second factor by $(n-m)!$ monomials of degree $n-m$. Hence (31) is the sum of $\binom{n}{m} \cdot m! \cdot (n-m)! = n!$ monomials of the form

$$\pm a^{p_1}_{s_1} \dots a^{p_m}_{s_m} a^{r_1}_{t_1} \dots a^{r_{n-m}}_{t_{n-m}} \qquad \dots \quad (32)$$

$u_1, \ldots, u_m$ is an arbitrary permutation of s, and $v_1, \ldots, v_{n-m}$ an arbitrary permutation of t_i. The sign + has to be taken if, and only if, both permutations are of the same kind, *i.e.*, if

$$u_1, \ldots, u_m, v_1, \ldots, v_{n-m} \qquad \ldots (33)$$

is an even permutation of $1, \ldots, n$,

and therefore also an even permutation of $p_1, \ldots, p_m, r_1, \ldots, r_{n-m}$.

Every permutation of $1, \ldots, n$ is a (33), the number of the monomials in (31) is $n!$, and therefore the sum becomes L. In the case of an odd permutation the sign of every monomial in (31) is changed, and therefore we get $-$L. When the numbers p and r are not all different, (31) becomes a determinant with two identical rows, and from § 10, 4 it follows, that it is vanishing.

The relation between the two factors of each term of (31) is a reciprocal one; they are called *Cofactors*. We can extend the notion of cofactor and the formula (31) also to the case $n=2$, by taking:

$$L_1^1 = a_1^1, \; L_2^2 = a_2^2, \; L_2^1 = a_2^1, \; L_1^2 = -a_1^2.$$

Often 14 is used in the case $m = n-1$, *i.e.*:

$$\sum_s a_s^j L_s^i = 0, \text{ for } i \neq j$$
$$= L, \text{ for } i = j. \qquad \ldots (34)$$

From 10 it follows that (34) holds also when the upper and the lower indices are interchanged; hence we get:

$$\Sigma a_i^s L_j^s = 0, \text{ for } i \neq j$$

$$= L, \text{ for } i = j. \qquad \ldots (34')$$

If an arbitrary Matrix D has the co-ordinates $[k, i] = a_k^i$ we will write:

$$D = ((a_k^i)).$$

On selecting some rows and some columns of D, we get new matrices

$$D_{k_1, \ldots, k_s}^{i_1, \ldots, i_s} = ((b_i^r)), \quad a_{k_r}^{i_r} = b_r^r.$$

If $u = v = m$, the matrices define determinants, called the *Minors* of D of order m

$$\det D^{i_1, \ldots, i_m}_{k_1, \ldots, k_m} = \det b^i_k . \qquad \ldots (35)$$

15. If $L^{p_1, \ldots, p_m}_{s_1, \ldots, s_m}$ and $L^{r_1, \ldots, r_{n-m}}_{t_1, \ldots, t_{n-m}}$ are cofactors, then

$$L^{p_1, \ldots, p_m}_{s_1, \ldots, s_m} = \det D^{r_1, \ldots, r_{n-m}}_{t_1, \ldots, t_{n-m}} .$$

Proof: The replacement (29), by which we get $L^{p_1, \ldots, p_m}_{s_1, \ldots, s_m}$ from L, annihilate in (25) every summand not having the factor $a^{p_1}_{s_1} \ldots a^{p_m}_{s_m}$, and in the remaining monomials this factor will be replaced by 1. the summands of $L^{p_1, \ldots, p_m}_{s_1, \ldots, s_m}$ are therefore just the summands of a determinant composed of the rows $(r_1), \ldots, (r_{n-m})$, and the columns $<t_1>, \ldots, <t_{n-m}>$. For getting the proper sign we must realise that the permutation of the $n-m$ rows is of the same character (even or odd) as the permutation of the $n-m$ columns if and only if the permutation of the original n rows is of the same character as the permutation of the original n columns. This condition is however satisfied by the definition of the co-factors.

16. If all minors of D of order m vanish, then the minors of higher order vanish also.

Proof: Using the formula (31) we can develop an arbitrary minor of order $m+f$ as a homogeneous linear function of degree m with co-efficients being minors of order f.

Definition 14: If there is a non-vanishing minor of D of order r but every minor of order $r+1$ vanishes, r is called the *Rank* of D.

17. The matrix D, the vector-space generated by its rows, and the vector-space generated by its columns have all the same rank.

Proof: If r is the rank of the vector-space of the row-vectors, there is a linear homogeneous relation between every set of $r+1$ row-vectors, and this relation holds also, if we cut away some of the columns. Hence from § 10, 4 it follows, that the minors of order $r+1$ are all vanishing, and

the rank of D is at most r. However if R is the rank of D every minor of order $>$ R vanishes, but a minor of order R is not vanishing. Without restriction of the generality we suppose $D^{1,\ldots,R}_{1,\ldots,R} \neq 0$. Let (w) be an arbitrary row.

In $D^{1,\ldots,R,w}_{1,\ldots,R,R+1}$ the co-factors of $a^1_{R+1},\ldots,a^R_{R+1},a^w_{R+1}$ will be called respectively $\qquad A_1,\ldots,A_R,\ A_w = D^{1\ldots R}_{1\ldots R} \neq 0.$

Hence the co-ordinates of $a = a^1A_1 + \ldots + a^RA_R + a^wA_w$ are either determinants with two equal rows, or minors of D of order R+1, and therefore vanishing. Hence $a = 0$, *i.e.*, every a^w depends on $a^1,\ldots,a^R$, hence the rank of the vector-space is at most R. As we have seen in the first part of the proof, the rank of the vector-space is not smaller than R, and therefore both ranks are equal. The rank of D is not changed, if the rows and the columns of D are interchanged, hence the rank of D equals also the rank of the vector-space generated by the columns of D.

§ 13. Elimination by Determinants.

Theorem XI. The equations (2) are solvable if and only if the matrices

$$M = \begin{pmatrix} a_1,\ldots,a_n \\ \ldots\ldots\ldots \\ k_1,\ldots,k_n \end{pmatrix} \quad \text{and} \quad \overline{M} = \begin{pmatrix} a_1,\ldots,a_n,a_0 \\ \ldots\ldots\ldots\ldots \\ k_1,\ldots,k_n,k_0 \end{pmatrix} \tag{36}$$

have the same rank.

Proof: From 18 it follows that the ranks of the matrices are the same as the ranks of the vector-spaces V and $\overline{V}$ of Theorem VI. Hence the Theorem XI follows from the Theorem VI.

In order to get the rank r of a matrix by the help of determinants, it is not necessary to calculate each minor; it is sufficient to state, that one minor of order r is not vanishing, and that all minors of order $r+1$ vanish. The rank does not change by row-addition, or by column-addition; it is often useful to simplify matrices by these operations.

Corollary: If the matrices M and $\overline{M}$ of (36) have the same rank, then the matrix formed by an arbitrary set of rows $(p_1),\ldots,(p_r)$ of M has the same rank as a matrix formed by the same rows of $\overline{M}$.

Proof: As the non-homogeneous equations belonging to M and $\overline{\mathrm{M}}$ are solvable, the rows (p_1) ….. (p_t) define solvable equations, and therefore the two matrices defined by these t rows have the same rank.

Theorem XII. If the matrices M and $\overline{\mathrm{M}}$ of Theorem XI have the same rank r, and a non-vanishing minor of M has the rows (p_1) ,…., (p_r), then the solutions of (2) are identical with the solutions of the equations belonging to these rows.

Proof: From the Corollary it follows that the rank of the matrix formed by the (p_1) ,…., (p_r) is also r. The solutions of the homogeneous equations defined by the r rows form a vector-space X of rank $n-r$, including all solutions of (2/H). As the rank of the vector-space of the solutions of (2/H) is also $n-r$, it follows from §4, 7 that every solution of the r linear homogeneous equations is also a solution of (2/H). Hence from Theorem I it follows that we get every solution of (2) by adding all vectors of X to an arbitrary solution of the r non-homogeneous equations, *i.e.*, the solutions of these equations are just the solutions of (2).

By the Theorems XI and XII the problem of solving a system of linear equations by elimination has been reduced to the following: Find out the solutions of

$$\begin{array}{l} a_1^1\, x_1 + \ldots + a_r^1\, x_r + a_{r+1}^1\, x_{r+1} + \ldots + a_n^1\, x_n = a_0^1 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_1^r\, x_1 + \ldots + a_r^r\, x_r + a_{r+1}^r\, x_{r+1} + \ldots + a_n^r\, x_n = a_0^r \end{array} \tag{37}$$

when the determinant det A of the matrix

$$\mathrm{A} = \begin{pmatrix} a_1^1 ,\ldots, a_r^1 \\ \cdots\cdots\cdots \\ a_1^r ,\ldots, a_r^r \end{pmatrix} \quad \text{is not vanishing.}$$

Theorem XIII. Let A_k^i be the co-factors of a_k^i in A, then we get the solutions of (37) by

$$\det \mathrm{A}\; x_k + \sum_i a_{r+1}^i \mathrm{A}_k^i x_{r+1} + \ldots + \sum_i a_n^i \mathrm{A}_k^i x_n = \sum_i a_0^i \mathrm{A}_k^i \tag{38}$$

$k = 1, \ldots, r$.

Proof: By multiplying the equations (37) respectively by the co-factors A_k^i (k being constant), and adding, we get the equations (38).

These equations are therefore necessary conditions for the solutions of (37). The rows of (38)—including the right sides of the equations—are dependent on the rows of (37). The rank of the matrix composed of the rows of (37) and (38) is therefore also r. The determinant formed by the first r columns of (38) is a power of det A and therefore $\neq 0$. Hence from Theorem XII it follows that the solutions of (38) are identical with the system of all solutions of (37) and (38), and therefore (38) is also a sufficient condition for the solutions of (37).

§ 14. Linear Transformations.

Let

$$A = \begin{pmatrix} a_1^1 & \dots & a_n^1 \\ \dots & \dots & \dots \\ a_1^n & \dots & a_n^n \end{pmatrix} \tag{39}$$

be a matrix with n rows and n columns.

The row-vectors are called $\alpha^1, \dots, \alpha^n$,

the column-vectors are called $\alpha_1, \dots, \alpha_n$.

We consider the equations

$$\sum_k a_k^i x_k = y_i \quad , i=1, \dots, n. \tag{40}$$

To every n-vector $\xi = (x_1, \dots, x_n)$ corresponds an n-vector $\eta = (y_1, \dots, y_n)$; (40) is called a *linear transformation*, and we will express it by

$$\xi \longrightarrow \eta \ .$$

Theorem XIV. By a linear transformation (40) the n-vectors ξ are transformed to the n-vectors of a vector-space, whose rank equals the rank of A.

Proof: If $\xi^1 \longrightarrow \eta^1$, $\xi^2 \longrightarrow \eta^2$, then $\xi^1 + \xi^2 \longrightarrow \eta^1 + \eta^2$

and $c\xi^1 \longrightarrow c\eta^1$, for every number c.

Hence from §4, 10 it follows, that η form a vector-space H. Two n-vectors ξ are transformed to the same n-vector η, if and only if the difference is transformed to 0, *i.e.*, if the difference belongs to the vector-space Z of the solutions of the homogeneous system belonging to (40). Using the methods of §4, 7, or §8, 3, we will find out a basis $\zeta^1, \dots, \zeta^{n-r}$, $\xi^1, \dots, \xi^r$ of the vector-space of ξ, so that the $n-r$ first n-vectors form

a basis of Z. If $\xi^i \longrightarrow \eta^i$, an arbitrary n-vector is transformed: $\Sigma c_i \xi^i + \Sigma d_k \zeta^k \longrightarrow \Sigma c_i \eta^i$, and from the definition of ζ^i it follows that this n-vector vanishes if and only if $c_1 = \ldots = c_r = 0$. Therefore $\eta^1, \ldots, \eta^r$ form a basis of H, and r is the rank of H.

Theorem XV. A representation of the n-vectors ξ by the n-vectors $f(\xi)$ with the properties: $f(\xi^1 + \xi^2) = f(\xi^1) + f(\xi^2)$, $f(c\xi) = cf(\xi)$, is a linear transformation.

Proof. Let $f(e^k) = \beta_k = \Sigma b^i_k e^i$, $\xi = \Sigma x_k e^k$, then $f(\xi) = \Sigma b^i_k x_k e^i$.

Hence the representation is a linear transformation with the matrix $((b^i_k))$.

If $y_i = \Sigma a^i_k x_k$, $x_k = \Sigma b^k_j z_j$, then $y_i = \Sigma a^i_k b^k_j z_j = \Sigma d^i_j z_j$, where

$$d^i_j = \Sigma a^i_k b^k_j = \mathrm{S}.\alpha^i \beta_j, \tag{41}$$

α^i being the row-vectors of $\mathrm{A} = ((a^i_k))$, and β_j the column-vectors of $\mathrm{B} = ((b^k_j))$. The matrix $\mathrm{D} = ((d^i_j))$ is called the product

$$\mathrm{D} = \mathrm{A} \cdot \mathrm{B} \tag{42}$$

If we consider 3 matrices A, B, C, and their products (A·B)·C and A·(B·C), we get in both cases a matrix with the co-ordinates $g^i_l = \sum_{k,j} a^i_k b^k_j c^j_l$; hence:

Theorem XVI. The associative law holds for the multiplication of matrices.

§15. Decomposition of Matrices.

A matrix $\mathrm{D} = ((d^i_k))$, $d^i_i = d_i$, $d^i_k = 0$, when $i \neq k$, is called a *Diagonal-matrix*. A matrix $\mathrm{E}_{rs}(\lambda) = ((e^i_k))$, $e^i_i = 1$, $e^r_s = \lambda$, and every other $e^i_k = 0$, is called an *Elementary-matrix* (defined only for $r \neq s$). If A has the row-vectors $\alpha^1, \ldots, \alpha^n$, and the column-vectors $\alpha_1, \ldots, \alpha_n$, then

D·A has the row-vectors $d_1\alpha^1, \ldots, d_n\alpha^n$

A·D has the column-vectors $d_1\alpha_1, \ldots, d_n\alpha_n$

$\mathrm{E}_{rs}(\lambda) \cdot \mathrm{A}$ has the row-vectors $\beta^i = \alpha^i$,

for $i \neq r$ $\quad \beta^r = \alpha^r + \lambda\alpha^s$ ✓ (43)

$\mathrm{A} \cdot \mathrm{E}_{rs}(\lambda)$ has the column-vectors $\gamma_i = \alpha_i$,

for $i \neq s$ $\quad \gamma_s = \alpha_s + \lambda\alpha_r$. ✓

Theorem XVII. An arbitrary matrix A of n rows and n columns is a product $A = P_1 \cdot D \cdot P_2$, where P_1 and P_2 are products of elementary-matrices, and D is a diagonal-matrix.

Proof. From (43) it follows that the theorem is identical with the following: A can be transformed to a diagonal-matrix by row-additions $a^r \longrightarrow a^r + \lambda a^s$, and by column-additions $a_r \longrightarrow a_r + \lambda a_s$. In order to prove this proposition, we use the method of "sweep-out" in a little modified manner. If every element of A vanishes, A is a diagonal-matrix; if all elements do not vanish, we can make $[1, 1] \neq 0$ by row-additions and column-additions of the type mentioned above, and by the same kind or operations we can sweep out the first row and the first column. On continuing this procedure we get a diagonal-matrix.

Theorem XVIII. The determinant of $A \cdot B$ is equal to the product of the determinants of A and B, *i.e.*,

$$\det (A \cdot B) = \det A \cdot \det B.$$

Proof. From (43) it follows, that the determinant of a matrix does not change, when the matrix is multiplied with an elementary-matrix, and therefore the determinant does not change, when the matrix is multiplied with a product of elementary-matrices. Hence, if $B = P_1 \cdot D \cdot P_2$, $\det B = \det D = d_1 \ldots d_n$. On the other hand $\det (A \cdot B) = \det (A \cdot P_1 \cdot D)$ holds. As we get $A \cdot P_1 \cdot D$ from $A \cdot P_1$ by multiplying the columns with $d_1 \ldots d_n$, we get: $\det (A \cdot B) = \det (A \cdot P_1 \cdot D) = \det (A \cdot P_1) \cdot d_1 \ldots d_n = \det A \det B$.